

Factors affecting the frequency of citation of an article

Rafael Repiso ¹, Alicia Moreno-Delgado ^{*1}, Ignacio Aguaded ²

How to cite (APA): Repiso, R., Moreno-Delgado, A., & Aguaded, I. (2020). Factors affecting the frequency of citation of an article. *Iberoamerican Journal of Science Measurement and Communication*; 1(1). <https://doi.org/10.47909/ijsmc.08>

Received date: 17 June 2020

Accepted date: 5 August 2020

Handling editor: Carlos Luis González-Valiente

Copyright: © 2020 Repiso et al. This is an open access article distributed under the terms of the CC BY-NC 4.0 license which permits copying and redistributing the material in any medium or format, adapting, transforming and building upon the material as long as the license terms are followed.

The quality of a scientific publication has long been associated with the number of citations it receives from other studies. In this sense, a citation is a hypertextual element resulting from one author's recognition of earlier research publications. The act of referring to a previous study is associated with the recognition of that study and an acknowledgment—regardless of whether the citation is positive or negative—that it has contributed to the development of new work currently in progress. In the field of communication and scientific evaluation, the relevance of citations is clear since they constitute a substantial part of most bibliometric indicators. The Impact Factor, H Index, or Crown Indicator, among others, are calculated from citations. Moreover, scientific search engines such as Google Scholar sort results by the number of citations (Rovira et al., 2018). Therefore, the citation is one of the major elements used when assessing the impact of research and—together with scientific production—it facilitates the aggregate-level study of knowledge generated by authors, journals, and institutions, among others. The databases that are most commonly used in the evaluation and study of science (Web of Science, Scopus or Dimensions) construct their impact, repercussion and similarity indicators from records of the principal descriptive elements of a publication (title, authors, subject matter, etc.) and from references to indexed studies. These databases do not store full texts; they only gather what is needed to identify the records and generate the

indicators.

In recent years, other means of analyzing the quality of published research (altmetrics, social networks, visualizations, or downloads) have emerged as a consequence of technological

1 Universidad Internacional de La Rioja, Spain. (Rafael Repiso: rafael.repiso@gmail.com; ORCID: 0000-0002-2803-7505), (* Corresponding author, email: aliciamorenodelgado@gmail.com, ORCID: 0000-0002-3425-061X).

2 Universidad de Huelva, Spain. (Email: aguaded@uhu.es, ORCID: 0000-0002-0229-1118).

development. However, the citation system still seems more rigorous as it is linked to the process of creating and publishing new research. The citation does not simply recognize the act of sharing or visualizing a piece of work but will in fact lead to more complex action in the production of further scientific research. Hence, manipulation within the citation system may appear more complex than any alteration caused by the emerging metrics.

WHY IS A SCIENTIFIC PAPER CITED?

Science is cumulative, which is why written records are made of the knowledge generated; these enhance the perspective and vision of subsequent researchers in order to validate new knowledge. Therefore, a paper is cited in recognition of its contribution: authors cite earlier papers because they have "reused" some of their components to create a new study within the scientific cycle of innovation and progress. This may constitute results, a methodology, a theoretical reflection, or, even, the rejection or correction of affirmations presented in the earlier study. So, studies are cited because their contents are relevant and that relevance usually lies in their scientific results—although, as we have already seen, other "excuses" can lead authors to cite research.

Those studies that have had the greatest influence on science are also those that are most frequently cited. Bradford's Law (Bookstein, 1994) tells us that papers published in a few journals (20%) are those that accumulate the highest number of citations (80%); this led to the Science Citation Index. Consequently, this indicator has long been used in attempts to predict Nobel prize-winners (Garfield & Malin, 1968). The problem with the system is that the community sometimes fails to identify the quality of highly innovative work, or does so late, as in the case of Mendel's Laws. Research studies like this are known as "Sleeping Beauties" (Van Raan, 2004).

WHAT BASIC FACTORS ARE NEEDED FOR A PAPER TO BE CITED?

The essential factor leading to a study being cited is that it should make a significant contribution to the advance of science; that is, the relevance of the research in itself. But other essential dimensions exist:

- (a) **Accessibility:** Can users access research content?
- (b) **Dissemination:** Can journals publicize and communicate the articles they publish?
- (c) **Scientific authority:** Does any given paper gain by being supported by the authors or publishers by virtue of their personal prestige drawn from readers' perceptions (brand) of them as academic authorities (thoroughness, honesty, reliability, etc.)?

Accessibility

If a paper is to be citable it must have been published—i.e. made accessible to the public—otherwise, it can only be cited by its authors and its reliability will be questioned, as it has not been scrutinized (validated) by the scientific community. The highest level of accessibility is that provided by open access journals, which allow documents to be consulted without subscription and they may even be stored in more than one place (scientific journal, authors' website, institutional repository, etc.).

Dissemination

In a digital world in which information is plentiful, scientific dissemination and selection are crucial. An open access journal can potentially reach a global audience, but only the best journals reach global audiences and most of them require users to subscribe. Scientific journals are media, not just reviewers and article repositories; therefore, the best journals, apart from having larger readerships, ensure that their work is also covered by other media, ranging from scientific forums to the mass media.

Scientific authority

Another essential aspect of article citation is the scientific authority of the person who creates it or the journal that evaluates and publishes it. Readers trust the quality and veracity of the studies they read when these come from authoritative sources; all the more so since nowadays, it is impossible to read everything given the disproportionate growth in the number of scientific studies. Scientific authority comes from following a path of serious, rigorous, and successful production. Authors who do not yet own that authority can "borrow" it from the journals that publish their work. In fact, interaction provides the system with feedback: journals of excellence provide new researchers with authority as they must meet strict requirements to achieve publication; and authors accumulate authority that reflects on the journals in which they publish.

Finally, in order not to be misleading, we cannot but stress the difficulty facing most open access journals. Potentially, they guarantee maximum accessibility to papers but they have little scientific dissemination beyond their own ambit and the generic social media; moreover, only a very small percentage have relevant scientific authority in the field. Subscriber journals, such as *Nature* or *Science*, have an unquestionable authority that is the result of their lengthy success-filled trajectories; this is why their contents are the most widely disseminated. Authors who publish in these journals know that it is highly likely that after publication they will be sought out for interviews by the mass media. The scientific relevance of this nature (dissemination + authority) is as yet unknown to open access journals. Therefore, authors prefer to publish in journals with academic authority and dissemination, even if accessibility is more limited. What is striking is that, despite being restricted to subscription-only access, these journals are read the most.

WHAT OTHER FACTORS INFLUENCE ARTICLE CITATION FREQUENCY?

We would like to present these factors from a pedagogical perspective since many studies that deal with the means of increasing citations—or even exhaustive reviews of them (Ebrahim et al., 2013)—are excessively technical. The distribution of citations by paper has been proved imbalanced, with a few papers receiving most of the citations, and thus confirming Lotka's law (Lotka, 1926). Another widely studied phenomenon is the timing of citations, with results showing that citation validity periods differ as a function of the field of research: citations are valid for longer in engineering than in biomedicine (Larivière et al., 2013), and many papers receive only belated ("Sleeping Beauty") or intermittent recognition (Costas et al., 2010). The aims of the present paper are to identify a number of factors associated with obtaining citations; to explain these; and, finally, to offer authors a number of hopefully useful suggestions.

Due to the lack of a time frame within which any given paper can be cited, newly published studies can only be evaluated by association with the impact and quality of the journal in which they are published. Associating the Impact Factor with a journal is known to be

inaccurate; even more so when the Impact Factor is not normally distributed and a few papers receive most of the citations. However, in the academic world, this approach is commonly used to study the uncertain present and future impact of scientific studies. While the impact of the environment is associated with individual impact, other predictive factors allow us to predict the number of citations a document may receive. For example:

- a. Prior production by the authors. Any given study is usually the result of the researcher(s)/author(s) typical line of work. So, authors need to cite their own prior research. However, researchers who produce little or are inconsistent in their lines of research will cite themselves scarcely. Journals are interested in highly productive researchers with coherent thematic lines of research.
- b. Structural context of the work. Papers resulting from large-scale research projects are often cited for the same project in subsequent papers. Any research project would be expected to produce a number of articles on a specific topic as one of its goals.
- c. Scientific trends. One context in which the topic and the expiry of the results converge is that of unique events that respond to extraordinary situations such as that provoked by the current COVID-19 crisis. These issues represent opportunities since research is essential, however, there is such a flood of scientific research that only a few studies that provide solutions and exhaustive analysis actually receive citations. A word of caution here: this type of "scientific trend" is of short-term interest and the studies they give rise to lose interest rapidly and soon cease to receive citations.
- d. Validity/Obsolescence (expiry) of results. Papers that analyze topics with a short expiry date (e.g., studies of monthly or annual data) tend to have less chance of being cited over time. For instance, citations of a study analyzing the 2015 PISA report's results will be concentrated in the years immediately after publication since this type of social reality is highly changeable. Hence it is better to analyze long time frames, rather than a single year (a decade in PISA results), enabling authors to present more meaningful periods.
- e. Quality of formal aspects. As we have indicated, papers are usually cited because of the originality of their contributions. However, other elements can be recognized and merit highlighting, and these may appear in the introduction, methodology, discussion, and conclusions or, even, in an original figure. So, we should not skimp on the quality and style of any section of a piece of work.
- f. Theoretical context of the study. The theoretical framework of a study is the context in which the research has been developed. Obviously, papers based on similar theoretical contexts are those most likely to cite each other. Hence, authors should aspire to locate their studies in an international context and connect the local with the global. A study that uses only Spanish theoretical referents will seem to have forgotten that research is being conducted elsewhere and, therefore, will find it difficult to be understood (and cited) beyond the borders of Spain.
- g. Types of work. Beyond content quality, several paper types are often cited more frequently: reviews and methodical papers can be easily replicated and obviously have genuine use.

SUGGESTED WAYS OF PRODUCING HIGHLY CITABLE PAPERS

1. Try to solve real, practical research problems, rather than increase ephemeral discussion in a field or create theoretical problems that lead nowhere. Citations are a recognition of the usefulness of research and of its contribution to the field.
2. Think big! If you are going to analyze something in a specific year, why not analyze it over a decade? This will become the gold standard for your chosen period of time. If you are going to study something in Spain, why not analyze it in a broader context (Latin America, Ibero-America, or Europe)? If help is needed, why not ask prominent authors to contribute?
3. Publish with researchers who contribute to and complement your knowledge and perspective. Publishing with researchers from other fields or countries (if they make a real contribution) both enriches your work and increases its potential target readership.
4. If your object of study is local or national, perhaps more readers (and therefore increased opportunities for citation) will come from national journals, which represent the paper's natural audience. Sometimes publishing on a national topic in a non-national journal forces you to sacrifice citations for the sake of journal impact, particularly if your paper needs to be translated into English. For example, a paper about Spanish scientific journals will reach more interested readers in journals such as *El Profesional de la Información* or the *Revista Española de Documentación Científica* than in *Scientometrics*.
5. If a publication is in English (or another language) and no extra effort is needed to produce a Spanish-language version because, say, a Spanish original was sent to the translator, we would strongly recommend you deposit the Spanish text as a preprint (after all, Spanish is the second most widely-spoken language in the world). Anything that makes things easier for readers will increase accessibility.
6. Avoid publishing in languages other than English or your own. If you are lucky enough to be able to use several languages, publish in the one that is most widely used.
7. Use a simple writing style and remember that both Ph.D. students and leading experts will be among your readers. This makes your work easier to read (and cite). Contextualize your results as simply as possible without losing sight of the fact that they are preceded by a vast volume of scientific literature and that this is international.
8. Contextualize your work comprehensively. Use many, varied references to improve the value of each paper. Each reference increases connectivity with other papers and potential readers, other than the authors of the papers themselves, may appear.
9. Pay close attention to the composition of the identifiable elements of the article. Firstly, these must enable the paper to be located by means of a careful selection of words in the title, abstract, and keywords; they must also encourage researchers to want access to the full text once they have located the manuscript.
10. Deposit versions in repositories and on scientific websites to increase others' chances of finding your manuscript. To do so, keep your scientific profiles (web page, research group, university, ORCID, Google Scholar, Publons, or Mendeley) up to date and give full details.
11. The estimated parameters of the repercussion of a research paper can be determined by

studying the citation background of other papers in the field—both your own and those of your co-authors—as well as those of the journal in which it will be published.

12. Finally, you must strictly adhere to ethical principles because current citation analysis is highly transparent and authors who act improperly are exposed to public scrutiny. Avoid self-citation if it is not really justified. You should not force others to cite your papers. It is utterly deplorable to see that many professors have Google Scholar H index ratings that are mainly based on citations made by their students in dissertations and essays. It is also sad to see researchers take advantage of their positions as reviewers to request that papers under evaluation cite their work.

Conflict of interest statement

The authors declare that there is no conflict of interest.

REFERENCES

- Bookstein, A. (1994). Towards a multi-disciplinary Bradford law. *Scientometrics*, 30(1), 353–361.
- Costas, R., van Leeuwen, T. N., & van Raan, A. F. J. (2010). Is Scientific Literature Subject to a 'Sell-By-Date'? A General Methodology to Analyze the 'Durability' of Scientific Documents. *Journal of the American Society for Information Science and Technology*, 61(2), 329–339. <https://doi.org/10.1002/asi.21244>
- Ebrahim, N. A., Salehi, H., Embi, M. A., Tanha, F. H., Gholizadeh, H., Motahar, S. M., & Ordi, A. (2013). Effective strategies for increasing citation frequency. *International Education Studies*, 6(11), 93–99. <https://doi.org/10.5539/ies.v6n11p93>
- Garfield, E., & Malin, M. V. (1968). Can Nobel Prize winners be predicted. 135th Meetings of the American Association for the Advancement of Science, Dallas, TX, 1–8.
- Larivière, V., Archambault, É., & Gingra. (2013). Long-Term Variations in the Aging of Scientific Literature: From Exponential Growth to Steady-State Science (1900–2004). *Journal of the American Society for Information Science and Technology*, 64(July), 1852–1863. <https://doi.org/10.1002/asi.20744>
- Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16(12), 317–323.
- Rovira, C., Guerrero-Solé, F., & Codina, L. (2018). Received citations as a main seo factor of google scholar results ranking. *Profesional de La Informacion*, 27(3), 559–569. <https://doi.org/10.3145/epi.2018.may.09>
- Van Raan, A. F. J. (2004). Sleeping Beauties in science. *Scientometrics*, 59(3), 467–472. <https://doi.org/10.1023/B:SCIE.0000018543.82441.f1>